

# 高機能化非単調ニューロンを用いた集積化神経回路の研究

著者	金城 光永
号	184
発行年	2000
URL	<a href="http://hdl.handle.net/10097/12874">http://hdl.handle.net/10097/12874</a>

氏名(本籍)	きんじょう みつなが 金城 光永 (沖縄県)
学位の種類	博士(情報科学)
学位記番号	情 博 第 184 号
学位授与年月日	平成 13 年 3 月 26 日
学位授与の要件	学位規則第 4 条第 1 項該当
研究科、専攻	東北大学大学院情報科学研究科 (博士課程) システム情報科学専攻
学位論文題目	高機能化非単調ニューロンを用いた集積化神経回路の研究
論文審査委員	(主査) 東北大学教授 中島 康治      東北大学教授 沢田 康次 東北大学教授 星宮 望 (工学研究科)

## 論文内容要旨

### 第 1 章 序論

時々刻々と変化する大容量情報を取り扱うにあたり、その処理の高速化のため、システムのハードウェア化は重要である。また、ハードウェア化することでより携帯性を高めた情報機器の実現が可能となる。その実現には、音声、画像及び映像などの高速認識が必要不可欠であり、これは、ニューラルネットワークの得意とするところである。そのニューラルネットワークの連想記憶問題において、従来の単調ではなく非単調な入出力特性を持つニューロンを用いることで、ニューラルネットワークの記憶容量が増加することが森田らにより報告 [1] され、数学的に解析及び証明された [1] [2] [3]。また、数値シミュレーションにより学習における非単調ニューロンの有効性も確認された [4] [5]。この非単調ニューロンはネットワーク規模はそのままに性能を向上させることが可能なため、その利用により高い機能性及び携帯性を備えたニューラルネットワーク情報処理システムの実現が期待できる。そこで本研究では、学習機能を持つニューラルネットワークに非単調ニューロンを利用し、アナログ技術を用いてそのニューロチップのハードウェア実現を目指す。

### 第 2 章 非単調ニューロンのハードウェア実現

非単調関数を実現する回路は図 1 で示されるゲイン可変台形波生成回路を用いて実現する。この台形波生成回路は台形波のゲインが調整可能で、非単調関数のパラメータ  $\theta$  に対応する閾値も自由に設定できる。この回路は前段の台形波生成部と後段の線形抵抗器 [6] からなる。前段はさらに上段と下段に分かれ、上段で台形波の立上りを形成し、また、上段と下段の減算で台形波の立ち下がり形成する。その前段部の出力電流  $I_{pre}$  は次式で表される。

$$I_{pre} = \begin{cases} I_{in} & 0 \leq I_{in} < I_0 \\ I_0 & I_0 \leq I_{in} < I_{th} \\ I_0 - I_{in} + I_{th} & I_{th} \leq I_{in} < I_0 + I_{th} \\ 0 & I_0 + I_{th} \leq I_{in} \end{cases} \quad (1)$$

ここで、 $I_0$  及び  $I_{th}$  はそれぞれバイアス電圧  $V_{GAIN}$  及び  $V_\theta$  によって制御されるバイアス電流である。この電流を後段の線形抵抗器によって電圧に変換し、非単調ニューロンの出力とする。

この区分線形型活性化関数回路を VDEC の ROHM 0.6 $\mu$ m CMOS プロセスで設計し実現した。そのチップ写真と、各制御電圧をパラメータ (measurement ( $V_c$ ,  $V_{GAIN}$ ,  $V_\theta$ )) とした入出力特性の測定結果をそれぞれ

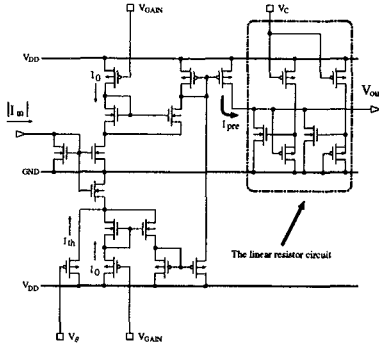


図 1: 非単調活性化関数回路



図 2: 非単調活性化関数回路のチップ写真:  $310\mu\text{m} \times 128\mu\text{m}$  (VDEC、ROHM、スタンダード CMOS、3 層 metal、2 層 Poly-Si、 $0.6\mu\text{m}$ )

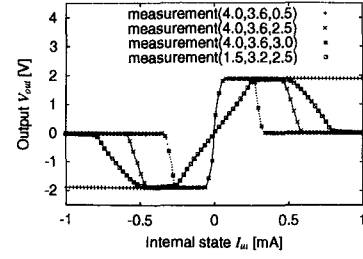


図 3: 非単調活性化関数回路の入出力特性

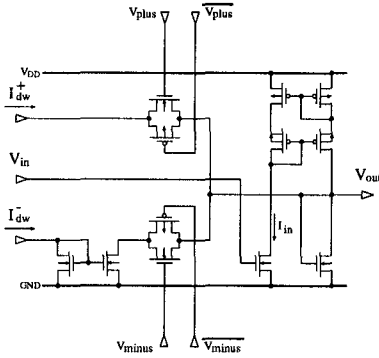


図 4: 荷重値更新回路



図 5: 荷重値更新回路のチップ写真:  $440\mu\text{m} \times 155\mu\text{m}$  (VDEC、ROHM、スタンダード CMOS、3 層 metal、2 層 Poly-Si、 $0.6\mu\text{m}$ )

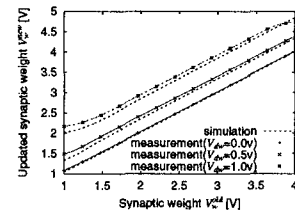


図 6: 荷重値更新回路の入出力特性

図 2 及び図 3 に示す。この結果より、制御電圧による閾値及びゲインパラメーターが調整可能な活性化関数回路が実現されたことが確認できる。

### 第 3 章 学習機能を持つシナプスのハードウェア実現

アナログ技術を用いた学習機能を持つシナプスのハードウェア実現において、アナログメモリの実現とその更新回路が重要な要素となる。本研究では、アナログメモリとして独自プロセスによる不揮発性アナログメモリ SDAM(Switched Diffusion Analog Memory) [7] を用いる。また、学習による荷重値更新に必要な線形性の高いアナログ電圧加減算器を図 4 により実現する。この荷重値更新回路は電流モードで構成されており、各素子が飽和領域で動作していると仮定すれば次式が成り立つ。

$$I_{W_{ij}} = \frac{\beta_n}{2} (V_{W_{ij}}^{(t)} - V_{th_n})^2 \quad (2)$$

$$V_{W_{ij}}^{(t+1)} = \sqrt{(V_{W_{ij}}^{(t)} - V_{th_n})^2 + \frac{2}{\beta_n} I_{\Delta W_{ij}}} + V_{th_n} \quad (3)$$

ここで、 $\beta_n$  及び  $V_{th_n}$  はそれぞれ  $n$  チャネル MOSFET のコンダクタンス及び閾値である。今、 $I_{\Delta W_{ij}}$  を次式

$$I_{\Delta W_{ij}} = \frac{\beta_n}{2} \{ V_{\Delta W_{ij}}^2 + 2(V_{W_{ij}}^{(t)} - V_{th_n}) V_{\Delta W_{ij}} \} \quad (4)$$

と設定する。ここで、 $V_{\Delta W_{ij}}$  は実際の荷重値更新値である。これより、式 (4) を式 (3) に代入すると、

$$V_{W_{ij}}^{(t+1)} = V_{W_{ij}}^{(t)} + V_{\Delta W_{ij}} \quad (5)$$

が得られ、荷重値が  $V_{\Delta W_{ij}}$  分線形に更新されることがわかる。ここで、式 (4) は、nMOS の線形領域でのドレイン・ソース間電流と類似しており、従って簡単な回路で実現できる。

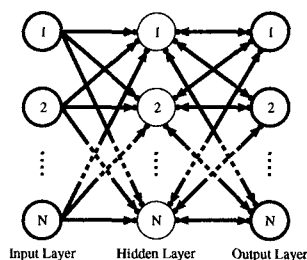


図 7: N-N-N DBM ネットワーク

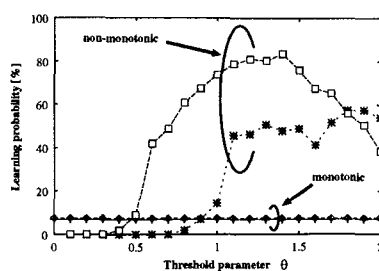


図 8: 4-4-1 DBM ネットワークにおける 4 パリティ問題の学習収束率

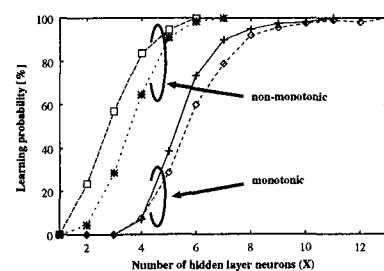


図 9: 隠れ層ニューロン数が異なるネットワークにおける 4 パリティ問題の学習収束率

この線形電圧加算器を VDEC の ROHM  $0.6\mu\text{m}$  ルール CMOS プロセスで設計し実現した。そのチップ写真及び入出力特性の測定結果をそれぞれ図 5 及び図 6 に示す。この結果より、 $1.5[\text{V}] \sim 3.5[\text{V}]$  の動作範囲内において、線形電圧加算が実現されたことが確認できる。

#### 第 4 章 非単調ニューロンネットワークのハードウェア実現

ニューラルネットワークの構成には、主に BP (BackPropagation) 学習及び DBM (Deterministic Boltzmann Machine) 学習 [8] を用いた方法が挙げられる。しかし、BP 学習では活性化関数の微分回路が必要となり、その回路は一般的に複雑であるため実現が困難である。さらに学習のために多数のローカルなメモリが必要であることから、ハードウェア化に不向きである。DBM 学習では、基本的に Hebb 則であり微分回路も必要無いため、大規模ニューロチップに適している。しかしながら、非単調ニューロンを用いる場合、BP 学習ではその収束性が保証されているが、DBM 学習では保証されていない。従って非単調ニューロンを用いた DBM の性能を解析する必要がある。そこで数値シミュレーションによりその性能を評価した。

数値シミュレーションに用いた決定論的ボルツマンマシンを図 7 に示す。このネットワークは、入力層、隠れ層及び出力層に  $N$  個のニューロンを有する。そのネットワークの隠れ層のニューロンに非単調ニューロンを適用し、単調ニューロンとの比較を行った。学習パターンの入出力対の値は  $\pm 1$  とし、そのため出力層のニューロンは単調ニューロンとした。また、学習の収束条件は、望ましい出力値とネットワークの出力値の誤差が 1% 未満となることとし、最大学習回数を 2000 回とした。図 8 に、単調ニューロンまたは非単調ニューロンを用いた 4-4-1DBM ネットワークに対し 4 パリティ問題を学習させた場合の学習収束率を示す。この結果より、非単調ニューロンを用いることで、学習の収束性が保証されていないにもかかわらず、従来のネットワークより高い学習性能が実現できたことがわかる。図 9 に隠れ層のニューロン数をパラメータとした学習収束率を示す。この結果より、少ないニューロン数で高い学習性能を引き出すことも可能で、実際、4 パリティ問題を 100% 学習可能な中間層の非単調ニューロン数は 6 個で、シナプスは 37 個である。これは単調ニューロンの場合の素子数と比較して約半分の規模であり、即ちネットワーク規模の削減が実現できることになる。

第 2 章及び第 3 章で述べた基本構成回路を利用して、非単調ニューロンを用いた 2-3-1DBM ネットワークのハードウェアを、東北大学電気通信研究所附属超高密度・高速知能システム施設で実現した。そのチップ写真を図 10 に示す。このチップサイズは  $5.2\text{mm} \times 5.2\text{mm}$  で、トランジスタ数は 2,360 個である。また、大規模集積化に向けて、図 11 に示すアナログメモリ無しの非単調ニューロンチップを VDEC ROHM  $0.6\mu\text{m}$  CMOS プロセスで設計し実現した。このチップサイズは  $4.5\text{mm} \times 4.5\text{mm}$  で、トランジスタ数は 2,036 個である。また、図 12 に非単調ニューロンチップでの 1 ニューロンにおける学習による荷重値更新と出力の変化を示す。この結果よりニューロチップ上で荷重値更新により正しく学習できたことを確認した。動作速度は測定系の問題で  $1\text{kHz}$  と低速だが、回路シミュレーターでは  $10\text{MHz}$  の想起動作及び  $1\text{MHz}$  の学習動作を確認した。

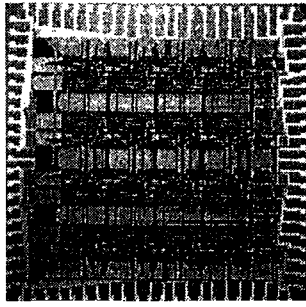


図 10: 2-3-1DBM ネットワークのチップ写真

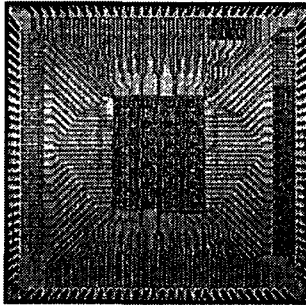


図 11: 非単調ニューロンチップのチップ写真

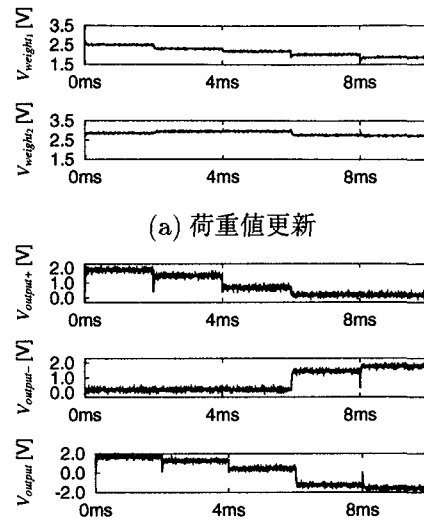


図 12: 1 ニューロンにおける学習による荷重値更新と出力の変化

## 第 5 章 結論

本研究では、ニューラルネットワークによる高速で携帯性の高い情報処理システムの実現に向けて、非単調ニューロンネットワークの基本構成要素のハードウェア実現及び学習性能の評価を行った。これより以下のことが結論できる。(1) 非単調ニューロンを用いることで、単調ニューロンネットワークと同等の性能を持つニューロチップが約半分の面積で実現できる。(2)  $0.35\mu\text{m}$  プロセス技術を用いれば、約  $20\text{mm}$  角のチップに 115 個のニューロンを含むニューロチップが実現でき、その CPS 及び CUPS はそれぞれ  $66.2 \times 10^9$ 、 $4.3 \times 10^9$  と見積もれる。

従って、非単調ニューロンを用いることで、従来のニューロチップを約半分の面積に削減でき、より携帯性の高い情報処理システムの実現が可能であると結論できる。

## 文献

- [1] M. Morita, Neural Networks, Vol.6, pp.115-126, 1993
- [2] M. Shiino and T. Fukai, J. Phys. A, vol.25, pp.L375-381, 1992.
- [3] H. Yanai and S. Amari, vol.3, pp.1385-1390, 1993.
- [4] M. Kinjo, S. Sato, and K. Nakajima, Proc. Int. Symp. NOLTA'98, vol.2, pp.455-458, 1998.
- [5] M. Kinjo, S. Sato, and K. Nakajima, Proc. IEEE Int. Conf. Neural Networks, no.2065, 1999.
- [6] M. Banu and Y. Tsividis, Electron. Lett., vol. 18, pp.678-679, 1982.
- [7] K. Nakajima, S. Sato, and T. Kitaura, IEICE Trans. Electron., vol.E78-C, pp.101-105, 1995.
- [8] G. E. Hinton, Neural Comp., vol.1, pp.143-150, 1989.

## 論文審査の結果の要旨

増大する高速な情報処理の要求に答えるため、ハードウェアの飛躍的な性能向上をベースにデジタルシステムの高性能化が推進されている。しかし一方ではノイマン型逐次直列処理の問題点などが指摘され、実時間処理やロバストな連想記憶などの知的処理の実現のために、逐次処理とは異なるシステムの検討も続けられている。そのなかで脳の構成的研究の一翼をも担うニューラルネットワークの構成については大規模な集積回路化が望まれており、応用への期待も大きい。筆者は、非単調ニューロンを採用することにより集積回路の規模を変えずにその性能を大幅に向上できることに着目し、また実世界とのインターフェースをより高速化するためにはアナログ情報処理が重要であるという観点から、アナログ技術を用いて学習機能を持つニューラルネットワークをハードウェア化するための基礎的研究を行った。本論文はその成果をまとめたもので、全編5章からなる。

第1章は序論である。

第2章では、非単調ニューロンの性能とその評価及びハードウェア実現について述べている。筆者はアニーリング法などの様々な手法や各種の問題にフレキシブルに対応できるように非単調ニューロンを設計し、アナログ技術を用いてハードウェア化を行った。

第3章では、学習機能を持つアナログシナプスのハードウェア実現について述べている。筆者は製作プロセス及び設計規則などの改良によって不揮発性アナログメモリの性能向上をはかり、信頼性の高い記憶機能を持つシナプスのハードウェア化を行った。さらに線形性の高い電圧加算器を提案し、すべてアナログ技術を用いて学習機能を実現し、これによりチップ動作を律速する学習演算を高速化した。これは集積化ニューラルネットワークの実用化に対する重要な基礎的成果である。

第4章では、前章で提案し集積化した非単調ニューロン及び学習機能を持つアナログシナプスを用いてニューラルネットワークを実現し、その測定と数値解析から非単調ニューロンを用いたネットワークの学習性能を詳細に評価している。その結果、非単調ニューロンを用いたネットワークが示す高い学習性能を見出し、また集積回路の動作速度及びチップ面積の評価から、従来の単調ニューロンを用いた場合と同等の性能が、約半分のチップ面積及び消費電力で実現可能であることを示した。これはニューラルネットワークの集積化に関する興味深い成果である。

第5章は結論である。

以上要するに本論文は、非単調ニューロンを用いた集積化ニューラルネットワークの実現に関する有用な基礎を与えたものであり、システム情報科学の発展に寄与するところが少なくない。

よって、本論文は博士（情報科学）の学位論文として合格と認める。